CHAPTER 18
BIOMEDICAL SIGNAL ANALYSIS

Jit Muthuswamy
Department of Bioengineering, Arizona State University, Tempe, Arizona

18.1 INTRODUCTION

Any signal transduced from a biological or medical source could be called a *biosignal*. The signal source could be at the molecular level, cell level, or a systemic or organ level. A wide variety of such signals are commonly encountered in the clinic, research laboratory, and sometimes even at home. Examples include the electrocardiogram (ECG), or electrical activity from the heart; speech signals; the electroencephalogram (EEG), or electrical activity from the brain; evoked potentials (EPs, i.e., auditory, visual, somatosensory, etc.), or electrical responses of the brain to specific peripheral stimulation; the electroneurogram, or field potentials from local regions in the brain; action potential signals from individual neurons or heart cells; the electromyogram (EMG), or electrical activity from the muscle; the electroretinogram from the eye; and so on.

Clinically, biomedical signals are primarily acquired for monitoring (detecting or estimating) specific pathological/physiological states for purposes of diagnosis and evaluating therapy. In some cases of basic research, they are also used for decoding and eventual modeling of specific biological systems. Furthermore, current technology allows the acquisition of multiple channels of these signals. This brings up additional signal-processing challenges to quantify physiologically meaningful interactions among these channels.

Goals of signal processing in all these cases usually are noise removal, accurate quantification of signal model and its components through analysis (system identification for modeling and control purposes), feature extraction for deciding function or dysfunction, and prediction of future pathological or functional events as in prosthetic devices for heart and brain. Typical biological applications may involve the use of signal-processing algorithms for more than one of these reasons. The monitored biological signal in most cases is considered an additive combination of signal and noise. Noise can be from instrumentation (sensors, amplifiers, filters, etc.), from electromagnetic interference (EMI), or in general, any signal that is asynchronous and uncorrelated with the

underlying physiology of interest. Therefore different situations warrant different assumptions for noise characteristics, which will eventually lead to an appropriate choice of signal-processing method.

The focus of this chapter is to help the biomedical engineer or the researcher choose the appropriate representation or analysis of the signal from the available models and then guide the engineer toward an optimal strategy for quantification. This chapter is not meant to be an exhaustive review of biosignals and techniques for analyzing. Only some of the fundamental signal-processing techniques that find wide application with biosignals are discussed in this chapter. It is not structured to suit the reader who has a scholarly interest in biomedical signal-processing techniques. For a more detailed overview of biomedical signal-processing techniques, the reader is referred to Refs. 1 and 2. This chapter will not deal with measurement issues of the signal. The reader is assumed to have acquired the signal reliably and is poised to make decisions based on the signals. This chapter will help the reader navigate his or her way from the point of signal acquisition to the point where it is useful for decision making.

A general classification of biomedical signals is attempted in Sec. 18.2. This will enable the reader (user) to place his or her signal of interest in the appropriate class. Subsequently, the sections are outlined according to different techniques for signal analysis. As far as possible, the first paragraph of each section generally outlines the class(es) of signals for which the corresponding technique is best suited. Toward the end of each section, appropriate MATLAB functions useful for analysis are indicated. Each section is then illustrated by an application.

18.2 CLASSIFICATIONS OF SIGNALS AND NOISE

The biomedical signal sources can be broadly classified into *continuous processes* and *discrete-time* or *point processes*. Each of these types of signals could be deterministic (or predictable), stochastic (or random), fractal, or chaotic. The continuous processes are typically encountered in one of the following situations.

18.2.1 Deterministic Signals in Noise

Examples of this type are ECG or single-fiber EMG signals in noise. The measured signal $x(t)$ can be represented as follows:

$$x(t) = s(t) + n(t) \qquad \dagger$$

where $s(t)$ is the actual deterministic signal and $n(t)$ is the additive noise. A segment of blood pressure shown in Fig. 18.1*a* is an example of a deterministic periodic signal. A Gaussian white noise assumption is valid in many biological cases. The goal in many situations is feature extraction under noisy (could be EMI, ambient, or instrumentation noise) conditions and subsequently correlating with the underlying physiological or pathological state.

18.2.2 Deterministic Signals (Synchronized to Another Stimulus Signal or Perturbation) in Noise

Examples of this type include all the different evoked responses (auditory, somatosensory, visual, etc.) and event-related potentials recorded in response to controlled stimuli administered to the body (or any biological system in general). These signals usually reveal functional characteristics of specific pathways in the body. For instance, evoked responses to peripheral somatosensory stimulation reveal the performance of the somatosensory pathway leading to the sensory cortex. A segment of cortical somatosensory evoked potential is shown in Fig. 18.1*b* that was obtained after averaging 100 stimulus-response pairs. Evoked responses or event-related potentials are usually

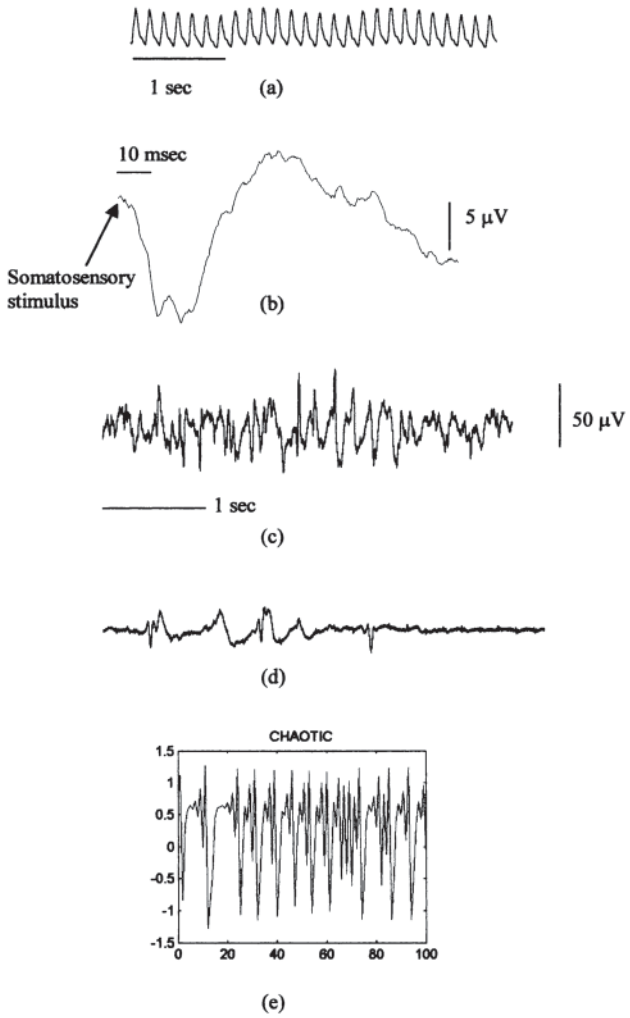

FIGURE 18.1 Types of biosignals: (*a*) *deterministic signal in noise*, illustrated by a segment of blood pressure signal recorded using a fluid-filled catheter in the femoral artery; (*b*) *deterministic signal (in noise) synchronized to an external cue or perturbation*, illustrated by an epoch of somatosensory evoked potential recorded from the somatosensory cortex in response to an electrical stimulus to the forelimb; (*c*) *stationary stochastic signal*, illustrated by a segment of EEG from the cortex that shows no specific morphology or shape, but the statistics of the signals are more or less stable or stationary in the absence of any physiological perturbations; (*d*) *nonstationary stochastic signal*, illustrated by a segment of EEG recorded from an adult rat recovering from a brain injury showing bursts of activity, but the statistics of this signal change with time; (*e*) *a chaotic signal* that was artificially generated resembles a stochastic signal but is actually generated by a deterministic dynamical system.

superimposed over spontaneous background electrical activity that is unrelated and hence asynchronous to the administered stimulation or perturbation. Therefore, signal-processing efforts have been directed toward extraction of evoked responses from the spontaneous background activity (could also be considered ìnoiseî in this case), noise (could be interfering signals such as ECG, EMG, or ambient noise) removal, and analysis of the evoked responses to quantify different components.

18.2.3 Stochastic or Random Signals

Examples of this type include EEGs, EMGs, field potentials from the brain, and R-R intervals from ECGs. Random signals lack the morphology of the signals found in the preceding two categories. Depending on the underlying physiology, the stochastic biosignals could be *stationary* (statistics of the signal do not change with time) or *nonstationary* (fluctuations in the signal statistics due to physiological perturbations such as drug infusion or pathology or recovery). A segment of EEG signal (random signal) that is stationary within the window of observation is shown in Fig. 18.1*c*, and an EEG signal that is nonstationary with alternate patterns of bursts and suppressions in amplitudes within the window of observation is shown in Fig. 18.1*d*. Signal-analysis techniques have been typically for noise removal and for accurate quantification and feature extraction.

18.2.4 Fractal Signals

Fractal signals and patterns in general are self-replicating, which means that they look similar at different levels of magnification. They are therefore scale-invariant. There is evidence to suggest that heart rate variability is fractal in nature. The branching of the airway into bronchioles seems to have a self-replicating nature that is characteristic of a fractal.

18.2.5 Chaotic Signals

Chaotic signals are neither periodic nor stochastic, which makes them very difficult to predict beyond a short time into the future. The difficulty in prediction is due to their extreme sensitivity to initial conditions, characteristic of these nonlinear systems. While fractal theory details the spatial characteristics of the nonlinear systems, chaos theory describes the temporal evolution of the system parameters or the dynamical variables. The essential problem in nonlinear biosignal analysis is to determine whether a given biosignal (a time series) is a deterministic signal from a dynamical system. Subsequently, signal analysis is usually done to determine the dimensionality of the signal and quantification of the dynamical states of the system. An example of a chaotic signal is shown in Fig. 18.1*e* that resembles a random signal but is actually generated by a deterministic dynamical system.

18.2.6 Multichannel Signals

Multichannel signals could include signals of any of the preceding five types but acquired using multichannel recording technology. Analysis and interpretation usually involve a matrix formulation of the single-channel analysis technique. A segment of four-channel multiunit activity (extracellular action potential signals) from thalamic and cortical structures is shown in Fig. 18.2. The goals of signal analysis are usually to identify correlation and hence synchrony among different channels, to achieve feature extraction under noisy conditions, and to identify underlying physiology.

The same six broad types of signals are found among point processes as well. The most common example of a point process is the action potential traces (recorded from either the extracellular or the intracellular space). Derived measures such as the neuronal firing rate histograms or cross-correllograms that are continuous in time are often used for analysis of these point processes.

The following sections are organized as follows. Section 18.3 deals with analysis of deterministic and stationary stochastic signals. Analysis of nonstationary signals is discussed in Sec. 18.4.

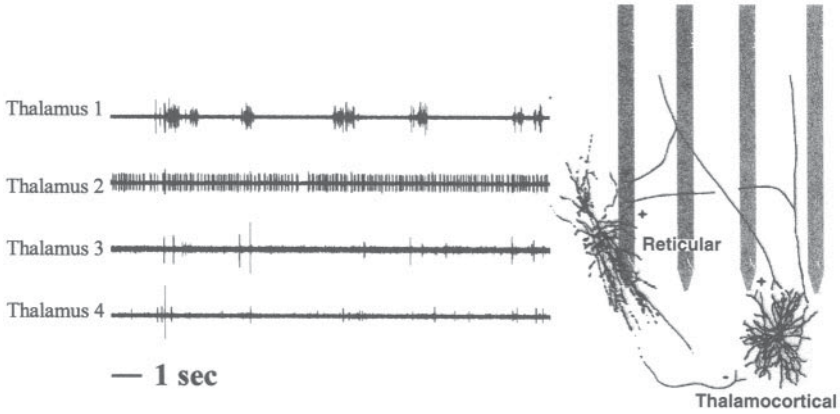

FIGURE 18.2 An example of a *multichannel biosignal* acquired using a multichannel microelectrode implanted in the thalamus and cortex of an adult rat. Each channel is a trace of multiunit activity (action potentials from a few neurons in the vicinity of each microelectrode as illustrated in the adjacent cartoon).

Subsequently, Sec. 18.5 deals with alternative orthogonal basis of representing biosignals that are optimal in certain physiological situations. Techniques for dealing with pairs of signals (both stochastic or a combination of stochastic and deterministic) are discussed in Sec. 18.6. Finally, Sec. 18.7 deals with analysis of fractal and chaotic signals. The statistics of the estimates are discussed wherever appropriate. By discussing these techniques separately, I am not suggesting that each of these techniques should be used in isolation. My hope is that the reader will get a deeper appreciation for each technique and will be driven toward creative solutions based on the data that could very well include a combination of several techniques.

18.3 SPECTRAL ANALYSIS OF DETERMINISTIC AND STATIONARY RANDOM SIGNALS

One of the most common analysis techniques that is used for biological signals is aimed at breaking down the signal into its different spectral (or frequency) components.

18.3.1 Fast Fourier Transforms

Signal Typesó Deterministic Biosignals (with or without Noise). Fast Fourier transform (FFT) is commonly used in analyzing the spectral content of any deterministic biosignal (with or without noise). I will discuss the issue of estimating the spectrum of the signal under noisy conditions in the following subsections. Discrete Fourier transform (DFT) allows the decomposition of discrete time signals into sinusoidal components whose frequencies are multiples of a fundamental frequency. The amplitudes and phases of the sinusoidal components can be estimated using the DFT and is represented mathematically as

$$X(k) = \frac{1}{N}\sum_{n=0}^{N-1} x(n)e^{-j(2\pi kn/NT)} \tag{18.1}$$

for a given biosignal $x(n)$ whose sampling period is T with N number of total samples (NT is therefore the total duration of the signal segment). The spectrum $X(k)$ is estimated at multiples of f_s/N, where f_s is the sampling frequency.

Fast Fourier transform (FFT) is an elegant numerical approach for quick computation of the DFT. However, users need to understand the resolution limitations (in relation to the signal length) and the effects of signal windowing on the accuracy of the estimated spectrum. In general, FFT does not work well for short-duration signals. The spectral resolution (or the spacing between ordinates of successive points in the spectrum) is directly proportional to the ratio of sampling frequency f_s of the signal to the total number of points N in the signal segment. Therefore, if we desire a resolution of approximately 1 Hz in the spectrum, then we need to use at least 1-s duration of the signal (number of points in 1-s segment = f_s) before we can compute the spectral estimates using FFT.

Direct application of FFT on the signal implicitly assumes that a rectangular window whose value is unity over the duration of the signal and zero at all other times multiplies the signal. However, multiplication in the time domain results in a convolution operation in the frequency domain between the spectrum of the rectangular window and the original spectrum of the signal $x(n)$. Since the spectrum of the rectangular window is a so-called Sinc function consisting of decaying sinusoidal ripples, the convolved spectrum $X(f)$ can be a distorted version of the original spectrum of $x(n)$. Specifically, spectral contents from one frequency component (usually the dominating spectral peak) tend to leak into neighboring frequency components due to the convolution operation. Therefore, it is often advisable to window the signals (particularly when one expects a dominating spectral peak adjacent to one with lower amplitude in the signal spectrum). Several standard windows, such as Hamming, Hanning, Kaiser, Blackman-Tukey, etc., are available in any modern signal-processing toolbox, each with its own advantages and disadvantages. For a thorough review on different windows and their effect on the estimation of spectral contents, the reader is referred to Ref. 3.

18.3.2 Periodogram Approach

Signal Typesó Deterministic Signals in Noise or Pure Stationary Random Signals. For deterministic signals in noise or pure random signals, Welchís periodogram approach can be used for estimating the power spectrum of the signal. The given length of the signal N is segmented into several (K) overlapping or nonoverlapping epochs each of length M. The power spectrum of each epoch is evaluated using FFT and averaged over K epochs to obtain the averaged periodogram. An implicit assumption in this method is that the statistical properties of the noise or the signal do not change over the length of the given sequence of data (assumption of *stationarity*.) Periodogram leads to statistically unbiased estimates (the mean value of the *estimate* of the power spectrum equals the true value of the power spectrum). Further, the variance (or uncertainty) in the individual estimates is inversely proportional to K. Therefore, it is desirable to increase the number of segments K in order to decrease the uncertainty in the estimates. However, increasing K will also decrease the resolution in the power spectra for reasons discussed in the preceding subsection. The power spectra using Welchís periodogram approach can be estimated using *spectrum* function in MATLAB. Since the power spectral values are estimates, one has to specify the confidence in these estimates for a complete description. When the epochs are nonoverlapping, the ratio of the power spectral estimate at each frequency to its actual value can be approximated to be a χ^2_{2K} random variable with $2K$ degrees of freedom.[4] Therefore, the interval within which the actual value is likely to lie can be estimated easily given the desired level of confidence. The function *spectrum* in MATLAB has options to estimate the confidence interval as well.

18.3.3 Parametric Methods

Signal Typesó Short-Duration Signals and Stationary Random Signals In Noise. When a high-resolution spectrum is desired and only a short-duration signal is available, parametric approaches outlined in this subsection provide a better alternative (as long as the parametric model is accurate) to FFT-based approaches for power spectral estimation in the case of *random or stochastic signals in noise*. For deterministic or periodic signals (with or without noise), Fourier transform-based approaches will still be preferable.[5] Examples of parametric modeling include the autoregressive (AR) model, the autoregressive moving-average (ARMA) model, and the moving-average (MA) model. So

far there is no automatic method to choose from AR, ARMA, or MA. The given signal is treated as the output of a linear time-invariant system (more advanced adaptive methods can be used to track time-varying model parameters) driven by a white Gaussian noise. The AR model can also be treated as an attempt to predict the current signal sample based on p past values of the signal weighted by constant coefficients. We estimate the best model by trying to minimize the mean squared error between the signal sample predicted by the model and the actual measured signal sample.

In an AR model, the signal $x(n)$ is represented in terms of its prior samples as follows:

$$x(n) = e(n) - a_1x(n-1) - a_2x(n-2) - \cdots - a_px(n-p) \tag{18.2}$$

where $e(n)$ = assumed to be zero mean white Gaussian noise with a variance of σ^2
p = order of the AR model
$x(n - i)$ = signal sample i time periods prior to the current sample at n
a_i = coefficients or parameters of the AR model

This representation can also be seen as a system model in which the given biosignal is assumed to be the output of a linear time-invariant system that is driven by a white noise input $e(n)$. The coefficients or parameters of the AR model a_i become the coefficients of the denominator polynomial in the transfer function of the system and therefore determine the locations of the poles of the system model. As long as the biosignal is stationary, the estimated model coefficients can be used to reconstruct any length of the signal sequence. Theoretically, therefore, power spectral estimates of any desired resolution can be obtained. The three main steps in this method are

1. Estimation of approximate model order (p)

2. Estimation of model coefficients (a_i)

3. Estimation of the power spectrum using the model coefficients or parameters

It is critical to estimate the right model order because this determines the number of poles in the model transfer function (between the white noise input and the signal output). If the model order is too small, then the power spectral estimate tends to be biased more toward the dominant peaks in the power spectrum. If the model order is larger than required, it often gives rise to spurious peaks in the power spectral estimate of the signal.

Several asymptotic model order selection methods use a form of generalized information criteria (GIC) that could be represented as $\text{GIC}(\alpha, p) = N \ln(\rho_p) + \alpha p$, where p is the model order, α is a constant, N is the number of data points, and ρ_p is the variance in the residual or error for model p.[6] The error or residual variance ρ_p can be determined using a (forward) prediction error $e_p(n)$ defined as

$$e_p(n) = \sum_{k=0}^{p} a_p(k)x(n-k), \qquad \dagger$$

where $a_p(k)$ is a parameter in the AR model of order p. The error variance is then simply a summation of squared forward prediction errors given by

$$\rho_p = \sum_{n=0}^{N-1} |e_p(n)|^2 \qquad \dagger$$

When the value of α is 2 in the expression for GIC, it takes the form of the Akaike information criterion (AIC).[7,8] The optimum model order p is one that minimizes the generalized or Akaike information criterion.[8] More recent methods to estimate model order for signal sequences with a finite number of sample points include predictive least squares (PLS) and finite sample information criteria (FSIC).[6]

Having determined the optimal model order for the given segment of signal, the model can be estimated using one of the following MATLAB functions: *arburg*, *arcov* (uses a covariance approach), *armcov* (uses a modified covariance approach), or *aryule* (uses the Yule-Walker equations

for estimation). The *arburg* function uses Burgís method[9] that minimizes both forward and backward prediction errors to arrive at the estimates of the AR model (similar to the modified covariance approach in *armcov*). The AIC model order estimation works well in combination with the Burgís method. Each of these functions uses different numerical procedures to arrive at the minimum mean square estimate of the AR model. The *aryule* function uses the autocorrelation estimates of the signal to determine the AR coefficients, which could potentially degrade the resolution of the spectrum. A brief discussion of the different AR parameter estimators and their application in ultrasonic tissue characterization is found in Ref. 10. Once the AR model has been estimated, the power spectra can be estimated by using the following expression:

$$P(f) = \frac{\sigma_p^2 T}{\left| 1 + \sum_{k=0}^{p} a_p(k) \exp(-j2\pi fkT) \right|^2} \tag{18.3}$$

where $P(f)$ = power spectral estimate at the frequency f
σ_p^2 = variance of the white noise input to the model
T = sampling period of the signal

The power spectrum of the signal can be estimated using the *pburg*, *pcov*, *pmcov*, or *pyulear* functions in MATLAB.

Statistics of the Spectral Estimates. The exact results for the statistics of the AR spectral estimator are not known. For large samples of stationary processes, the spectral estimates have approximately a Gaussian probability density function and are asymptotically unbiased and consistent estimates of the power spectral density. The variance of the estimate is given by

$$\text{Var}[P(f)] = \begin{cases} \frac{4p}{N} P^2(f) & f = 0 \text{ and } \frac{f_s}{2} \\ \frac{2p}{N} P^2(f) & \text{otherwise} \end{cases} \tag{18.4}$$

where f_s = sampling frequency of the signal
N = number of samples
p = model order[4]

In general, parametric methods are preferred for stochastic signals, provided the estimations of model order and the model parameters are done carefully.

APPLICATION *An 8000-point EEG signal segment sampled at 278 Hz is shown in the Fig. 18.3a. The power spectrum of the signal as evaluated using Welchís periodogram method is shown in Fig. 18.3b. The MATLAB function* spectrum *was used with a Hanning window of length 512, allowing for 256-point overlap between successive segments. The resolution of the spectrum is the ratio of the sampling frequency to the number of points, which turns out to be 0.54 Hz. The corresponding 95 percent confidence interval shown (upper curve) along with the power spectral estimate (lower curve) indicates the slightly higher uncertainty in the estimates around 5 Hz. An AR model order of 40 was chosen using the AIC shown in Fig. 18.3c. The power spectral estimate from the AR model using Burgís method in Fig. 18.3d shows a much higher resolution.*

18.4 SPECTRAL ANALYSIS OF NONSTATIONARY SIGNALS

A signal is nonstationary when the statistics of the signal (mean, variance, and higher-order statistics) change with time. The traditional spectral estimation methods just outlined will only give an averaged

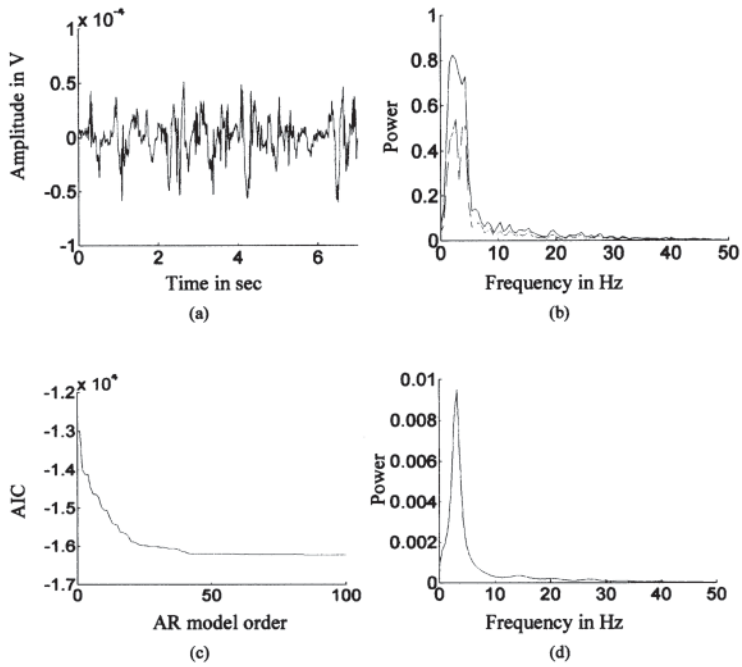

FIGURE 18.3 Comparison of a periodogram-based approach and a parametric approach to spectral analysis of stochastic signals, (*a*) A typical segment of EEG. (*b*) The power spectrum estimated using the periodogram approach (lower curve) along with the 95 percent confidence interval (upper curve). There is a higher uncertainty in the estimates in the low-frequency regime, (*c*) The autoregressive (AR) model order is estimated using an AIC criteria. The value of AIC seems to reach a minimum at approximately a model order of 50. (*d*) The power spectrum of the EEG signal estimated using an AR model of order 50. The power spectrum from the AR model has a higher resolution than (*b*).

estimate in these cases and will fail to capture the dynamics of the underlying generators. Alternative analysis techniques that determine time-localized spectral estimates can be used when the user suspects that the signal sequence under consideration is not under steady-state physiological conditions. These methods work well with both deterministic and random signals in noise.

Some of the commonly used algorithms for time-frequency representations of spectral estimates include the short-time Fourier transform (STFT), the Wigner-Ville transform, wavelet transforms, etc. A good summary of the mathematical basis of each of these techniques can be found in Ref. 11.

18.4.1 Short-Time Fourier Transform

The STFT as introduced by Gabor involves multiplying the signal by a short-duration time window that is centered around the time instant of interest. The Fourier transform of the product then gives an estimate of the spectral content of the signal at that time instant. The short-time-duration window is subsequently slid along the time axis to cover the entire duration of the signal and to obtain an estimate of the spectral content of the signal at every time instant. The signal is assumed to be stationary within the short-duration window. Mathematically,

$$P(m, n) = \int_{-\infty}^{\infty} e^{imt\omega_0} h(t - n\tau_0) x(t) dt \tag{18.5}$$

where $P(m, n)$ = short time Fourier transform (STFT) at time instant $n\tau_0$ and frequency $m\omega_0$
τ_0 and ω_0 = sampling interval and fundamental angular frequency, respectively
$h(t)$ = short-duration window
$x(t)$ = given signal

The STFT thus results in a spectrum that depends on the time instant to which the window is shifted. The choice of Gaussian functions for the short-duration window gives excellent localization properties despite the fact that the functions are not limited in time. Alternatively, STFT can also be viewed as filtering the signal ìat all timesî using a bandpass filter centered around a given frequency f whose impulse response is the Fourier transform of the short-duration window modulated to that frequency. However, the duration and bandwidth of the window remain the same for all frequencies.

18.4.2 Wavelet Transforms

The wavelet transform, in contrast, is a signal-decomposition (or analysis) method on a set of orthogonal basis functions obtained by dilations, contractions, and shifts of a prototype wavelet. Wavelets have been used extensively for processing biomedical images[12ñ15] and for processing almost every kind of biosignal with nonstationarities.[16ñ25] The main distinction between Fourier transform-based methods such as the STFT and wavelet transforms is that the former use windows of constant width, whereas the latter use windows that are frequency-dependent.[7,26] Wavelet transforms enable arbitrarily good time resolution for high-frequency components and arbitrarily good frequency resolution for low-frequency components by using windows that are narrow for high-frequency components and broad for low-frequency components. In general, the continuous wavelet transform can be represented mathematically as

$$S(a, b) = \int x(t)\psi^*_{a,b}(t)\, dt \tag{18.6}$$

where $x(t)$ = given signal
a = scale factor
b = time
$*$ = complex conjugate

The orthogonal basis functions denoted by $\psi_{a,b}(t)$ are obtained by scaling and shifting a prototype wavelet function $\psi(t)$ (also sometimes called a *mother wavelet*) by scale a and time b, respectively, as shown below:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right) \tag{18.7}$$

By adjusting the scale factor, the window duration can be arbitrarily changed for different frequencies. By choosing a and b appropriately, a discrete time version of the orthogonal basis functions can be represented as follows:

$$\psi(m, n) = a_0^{-m/2}\psi(a_0^{-m}t - nb_0) \tag{18.8}$$

where a_0 and b_0 are fundamental scale factor and time shift, respectively (m and n represent multiples of scale factor and the time shift).

Wavelet analysis works especially well with signals that have short durations of high-frequency components and long durations of low-frequency components, e.g., EEG signals or signals or variations in interbeat (R-R) intervals, etc. I describe in this subsection a simple procedure for

obtaining time-frequency representation of signals using wavelet transforms as outlined in Refs. 7, 27, and 28. A MATLAB function that implements the same procedure is detailed next.

1. Using FFT, find the Fourier transform $X(\omega)$ of the discrete time signal N-point signal $x(n)$.
2. For every equally spaced discrete frequency of interest f_i (>0), ($f_{\min} < f_i < f_{\max}$),
   - ***a***. Determine the scale factor $a_i = 1/f_i$.
   - ***b***. Denoting the analyzing wavelet by $\psi(n)$ (being implicitly dependent on the scale factor a) and its DFT by $\Psi(\omega)$, which is centered around ω_0, evaluate the product vector $Z(\omega) = \Psi(a\omega)X(\omega)$ for $\omega = -N/2$ to $N/2 - 1$.
   - ***c***. Determine the inverse Fourier transform of $Z(\omega)$ and scale the resulting time series by $\sqrt{a}$.

$$S(n, \omega) = \sqrt{a} \int_{\omega_{\min}/a}^{\omega_{\max}/a} \Psi(a\omega)X(\omega)e^{jn\omega}\, d\omega \tag{18.9}$$

   - ***d***. Repeat steps a to c for every discrete frequency within the range $f_{\min}$ and $f_{\max}$.
   - ***e***. The resulting output $S(n, \omega)$ is the time-frequency representation of the signal $x(n)$.

The Fourier transform of one of the prototype wavelets (called a *Morlet wavelet*) is given by

$$\Psi(\omega) = \exp[-(\omega - \omega_0)^2/2] + \text{small correction terms}\dagger$$

The correction terms are theoretically necessary to make the negative-frequency components of the term $\exp[-(\omega - \omega_0)^2/2]$ zero (an additional constraint imposed by Kronland-Martinet et al.[27]). In order to make the correction terms negligible, the center frequency ω_0 is chosen to be anywhere between 5.0 and 6.0. The corresponding time-domain form of the analyzing wavelet is a modulated Gaussian of the form (up to a constant normalization factor)

$$\psi(n) = e^{i\omega_0 n} \exp(-n^2/2)\dagger$$

The square of $S(n, \omega)$ is often referred to as a *scalogram*, and it gives an estimate of the time-localized power spectrum. A MATLAB function that can be used to implement this procedure is detailed below[28]:

```
function S=wt_morl(x,fs,nfft,nout,f0,fmin,fmax,pas,pass,sigm);
%function S=wt_morl(x,fs,nfft,nout,f0,fmin,fmax,pas,pass,sigm);
% time frequency analysis using Morlet wavelet
% x - signal vector
%fs: sampling frequency
%nfft: number of points in the signal
%nout: number of points not use in the beginning of signal ( =0)
%f0: center frequency of the wavelet = 1
%fmin: minimum frequency of the representation
%fmax: maximum frequency
%pas: increment in the scale between fmin and fmax
%pass: increment in time of the result (steps in time resolution)
%sigm: time frequency resolution general case = 0.7;
%S: result (it is better to plot abs(S))

y=fft(x(1 +nout:nfft+nout));
for f=fmin:pas:fmax
a = 1/f;
wavelet=exp(-2/sigm*(pi)^2*(([0:nfft/2-1,-nfft/2:-1]*(fs*a/nfft))-(f0)).^2);
temp=wavelet.í.*y;
t=ifft(temp).í;
S(fix(1/pas*(f-fmin))+1,:) = (a^0.5).*t(1:pass:nfft);
end;
```

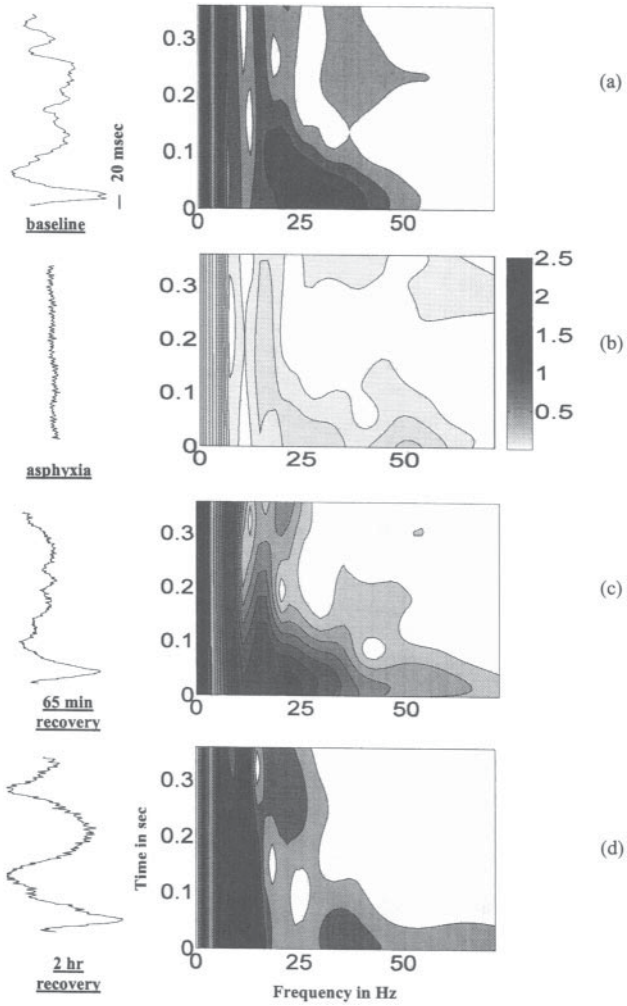

FIGURE 18.4 Analysis of a *nonstationary stochastic biosignal* using a wavelet-based approach for time-frequency localization of signal power. Somatosensory evoked potentials in response to an electrical stimulation to the contralateral forelimb recorded at different stages of brain injury are shown to the left of the vertical axis of the colored contour plots. The signals are recorded (*a*) at baselineó before the rat was subject to a controlled duration of asphyxia or oxygen deprivation; (*b*) during asphyxiaóshowing electrical silence or isoelectric behavior; (*c*) after approximately 65 minutes into recovery from the injuryó showing partial recovery of low frequency amplitudes; (*d*) after approximately 2 hours into recoveryóshowing greater recovery of the low-frequency and high-frequency generators in the evoked potentials. The time-frequency distribution of evoked potentials seems to indicate a greater vulnerability of the high-frequency generators in the brain to oxygen deprivation.

APPLICATION *I have used wavelet transforms to track the changes in the spectral components of the somatosensory evoked potentials with an ischemic injury (induced by asphyxia) to the brain in a rodent model, as illustrated in Fig. 18.4. The high-frequency components disappear first at the induction of injury and recover long after recovery of the low-frequency components after resuscitation of the animal. The results seem to suggest differential vulnerabilities in the generators of brain potentials. The somatosensory responses were obtained by averaging 100 stimulus-response pairs.*

18.5 PRINCIPAL COMPONENTS ANALYSIS

Signal Types— Deterministic or Stochastic Signals in Colored Noise. So far we have represented (by projection) biosignals in terms of Fourier basis functions, autoregressive parameters, or wavelet-basis functions depending on application. In general, the signal can be represented in terms of any convenient set of orthogonal basis functions. One such orthogonal basis, called the *principal components* of the given signal, has been widely used for analysis of biosignals. It is also called *Karhoenen-Loeve analysis* for signals with zero mean.

Broadly speaking, principal component analysis has been used for at least two different applications where signal representation using other basis functions does not provide the optimal solution. First, when the presence of a known signal waveform (hence deterministic) has to be detected in colored (implying correlation among noise terms) Gaussian noise, projecting the noisy signal along its principal components allows for optimal detection of signal. When the noise is white (independent, identically distributed) Gaussian, a matched filter (or a correlator, to be discussed in a subsequent subsection on cross-correlation) provides the most optimal detection. The second broad application of principal components analysis has been in pattern classification, where it has been used for identifying the feature vectors that are most discriminating. Selecting the most discriminating features leads to a reduction in the dimensionality of input feature vectors. Indeed, one could reasonably argue that the first application is a special case of the second. Nevertheless, I will treat them distinctly in this section and outline methods for both applications.

18.5.1 Detecting Deterministic or Known Signals in Colored Gaussian Noise

The observation vectors (successive time samples could also be used to construct the observation vector) are the sum of a known signal and noise terms that are correlated. The key idea in this application is to emphasize the projection of the signal along directions where the noise energy is minimal. The assumption is that we know the covariance matrix of the noise a priori. Even if it is not known, it can be easily estimated from the covariance of a *training set* (trial observations) of observation vectors. The element K_{ij} of the covariance matrix K is the estimated covariance between ith and jth components of the observation vector. Having estimated K, we find the eigenvalues and the corresponding eigenvectors of the matrix. If the matrix is positive definite, all the eigenvalues will be distinct, and the corresponding eigenvectors will be orthogonal. By suitably scaling the eigenvectors, we obtain a set of orthonormal basis functions ϕ_i, $i = 1, 2, \ldots, N$, where N is the dimensionality of the observation vector. The known signal vector $\mathbf{S}$ and the observation vector $\mathbf{X}$ are projected onto the set of orthonormal eigenvectors. Mathematically, the $(N \times N)$ matrix

$$\mathbf{\Phi} = \begin{bmatrix} \phi_1^T \\ \phi_2^T \\ \vdots \\ \phi_N^T \end{bmatrix}$$

represents the set of N orthonormal eigenvectors (each of dimension $N \lozenge 1$). $\mathbf{S}' = \mathbf{\Phi S}$ represents the projection of the known signal vector along each eigenvector. Similarly $\mathbf{X}'$ represents the projection of the received observation vectors (known signal plus colored noise) along the orthonormal eigenvectors. The optimal detector is designed by evaluating

$$\sum_{i=1}^{N} \frac{S_i'}{\lambda_i} X_i'$$

for every observation vector, where λ_i are the corresponding eigenvalues. This ratio (also called the *likelihood ratio* after normalization by the norm of $\mathbf{S}'$; it is really a cross-correlation of lag zero between projections of the known signal and projections of the observation vector) is compared against a threshold that is calculated from a priori known probabilities to detect the presence of the known signal. As the ratio indicates, the signal projection is emphasized (amplified) in directions where the eigenvalues are small, implying smaller noise energy in that direction. The threshold $?_0$ for an optimal Bayes detector is given by Ref. 29,

$$\Lambda_0 = \frac{p_0(C_{10} - C_{00})}{p_1(C_{01} - C_{11})} \tag{18.10}$$

where p_1 = a known a priori probability of the presence of the known signal (the frequency of occurrence of the known signal can be estimated by prior observation or knowledge of the generating process)
p_0 = $(1 - p_1)$
$C_{ij}(0 \leq C_{ij} \leq 1)$ = cost of deciding i when j actually occurred (where 0 indicates absence and 1 indicates presence of known signal in the noise)

Typically, C_{00} and C_{11} can be assumed to be zero, whereas C_{01} and C_{10} can each be assumed to be unity. However, sometimes the cost of false-positive C_{10} (detecting a signal when there is none) may be different from the cost of false-negative C_{01} (missing the known signal when it was actually present) depending on the application. Therefore, when the likelihood ratio exceeds the Bayes threshold, the detector indicates the presence of the known signal, and vice versa.

In the statistical analysis toolbox of MATLAB, functions *princomp* and *pcacov* enable the computation of principal components. The same method can be extended to detecting a finite number M of known signals. For further information on M-ary detection, the reader is referred to Ref. 29.

18.5.2 Principal Components Analysis for General Pattern Classification

Biosignals are often examined for recognizable patterns that indicate change in pathology or underlying physiology and hence aid diagnosis. In the absence of any signal analysis, the sample points in the measured biosignals have to be directly used to discriminate changes in underlying physiology. Analysis of the signal along spectral or wavelet components may not be the most discriminating since the components are not optimized for *discrimination* but are optimized for *spectral decomposition* and localization of energy. However, principal components analysis offers an elegant way to choose a set of orthonormal basis functions that are optimized specifically for maximum discrimination between two or more known classes of biosignals. In most cases, this method results in a drastic reduction in the number of signal components that one has to work with and also improves the classification or diagnosis as well.

Given a set of observation vectors (containing samples of the biosignals) from two different physiological or pathological states, we form a scatter matrix, which is really a scaled version of the sample covariance matrix. The scatter matrix $\mathbf{K}$ is defined as[30]

$$\mathbf{K} = \sum_{k=1}^{L} (\mathbf{X}_k - \mathbf{M})(\mathbf{X}_k - \mathbf{M})' \tag{18.11}$$

where L = number of N-dimensional observation vectors $\mathbf{X}_i$
$\mathbf{M}$ = sample mean of the observation vectors
t = transpose operator

All the observations from different physiological states are pooled together to form the scatter matrix. Since the scatter matrix is real and positive, all the eigenvectors of $\mathbf{K}$ are orthogonal. The eigenvectors are then normalized as in the preceding application to yield an orthonormal set of basis vectors. The magnitude of the corresponding eigenvalues indicates the degree of scatter along those eigenvectors. We then choose the eigenvectors corresponding to the largest eigenvalues as the preferred directions of projections and determine the projections of the subsequent observations along those directions.

APPLICATION *Principal component analysis is commonly used for sorting neuronal spikes (action potentials). Extracellular data collected using microelectrodes implanted in the brain typically record action potentials or spikes from more than one neuron. Different neurons typically register action potentials of different shapes and sizes at the microelectrode. A trace of multineuronal action potential data (or multiunit data) is shown in Fig. 18.5b. The three different action potentials found in the trace are shown separately in an enlarged fashion in Fig. 18.5a. The result of principal components analysis of the three action potential waveforms is shown in Fig. 18.5c. When plotted in the space spanned by the two most dominant principal components, the corresponding projections of the three different action potentials show three distinct clusters that are easily separable. Instead of using all the time samples that constitute the waveform of the action potentials, principal components analysis identifies just two features (projections along two dominant principal components or eigenvectors) that are sufficient for discrimination.*

18.5.3 Independent Components Analysis

Signal Typesó Stochastic Signals in Noise and Mixtures of Several Stochastic Signals. A recently developed extension of principal components analysis called *independent components analysis* (ICA) was originally used for *blind-source separation.*[31] In other words, when the measured signal vector (the number of channels of measured signal is given by the dimension of the vector) is an unknown linear mixture or combination of an equal number of independent, non-Gaussian sources, then this technique is useful in arriving at an estimate of the unknown original sources of the signal. For instance, when multiple channels of EEG are collected from several locations on the scalp, the potentials are presumably generated by mixing some underlying components of brain activity. While principal components analysis is useful in identifying the orthogonal directions (eigenvectors) that contain significant signal energy, independent components analysis is useful in identifying the independent, non-Gaussian components of the signal. This technique will not work if there is a strong indication that the original independent sources are Gaussian (if just one of the independent components is Gaussian, the ICA model may still be estimated[31]).

If the independent signal sources are represented by $\mathbf{s} = \{s_1, s_2, \ldots, s_N\}$ (could be N different kinds of sources) and an unknown *mixing* matrix $\mathbf{A}$ (inherent to the generating source or the medium between the source and the measurement point) generates the measured or received signals $\mathbf{r} = \{r_1, r_2, \ldots, r_n\}$, then the ICA signal model can be represented as $\mathbf{r} = \mathbf{A}\mathbf{s}$. This technique has been used successfully on EEG signals,[32,33] to decompose evoked potentials,[34] and to remove artifacts in magnetoencephalography (MEG).[35]

The goal of ICA is to design a matrix $\mathbf{F}$ specifying linear spatial filters that inverts the mixing process due to the $\mathbf{A}$ matrix. The matrix $\mathbf{F}$ is therefore often referred to as the *separation matrix* that helps to generate a scaled and permuted version $\mathbf{y} = \{y_1, y_2, \ldots, y_N\} = \mathbf{F}\mathbf{r}$ of the original sources $\mathbf{s}$. The basic principle of ICA is centered on the central limit theorem in statistics, which states that the distribution of a sum of independent random variables tends toward a Gaussian distribution under certain conditions. Therefore, the signal $\mathbf{r}$ (mixture of several independent non-Gaussian sources) will be more Gaussian. The idea then is to formulate a measure of non-Gaussianity and then project

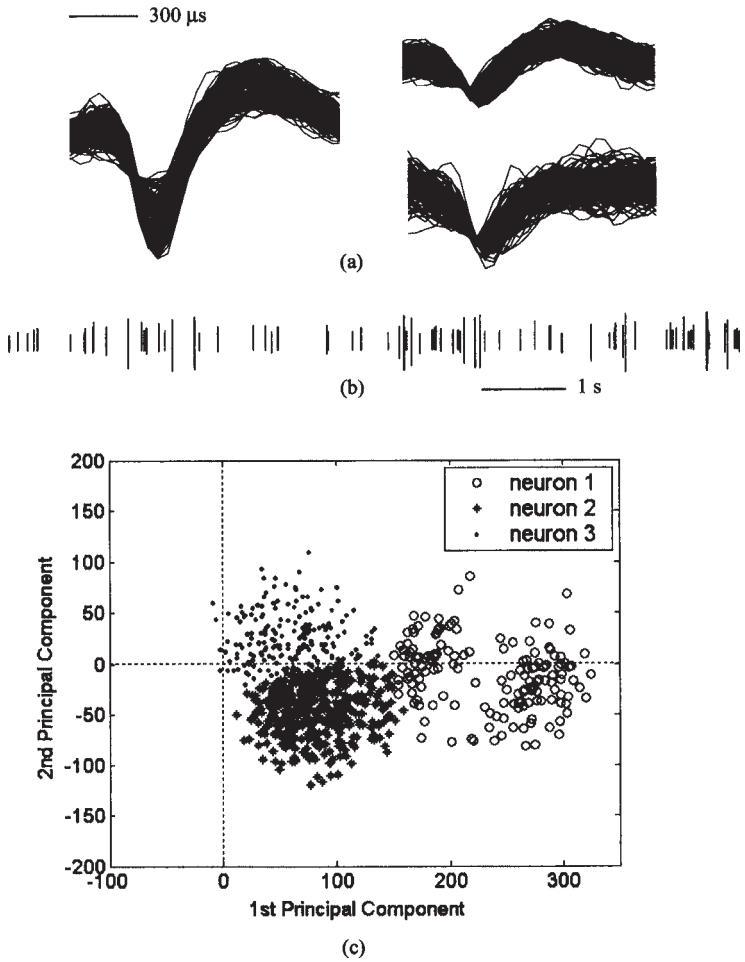

FIGURE 18.5 The measured biosignal (multiunit activity from a single microelectrode) is projected along an alternate (as opposed to Fourier basis functions shown earlier) set of orthogonal basis vectors (eigenvectors or principal components) for the purpose of *sorting* action potentials belonging to different neurons. The three different action potentials recorded by the microelectrode are shown in (*a*). Multiple traces within each action potential indicate multiple occurrences of similar action potentials superimposed to demonstrate variability. The three action potentials differ from each other in duration and peak amplitudes. On a more compressed scale in (*b*) the three action potentials appear as spikes of different heights. Plotting the action potential waveforms along two of their principal components in (*c*) gives a clear separation between the three different action potentials. Thus, instead of using all the points of the action potential waveform for sorting or classification, the use of projections along the principal components reduces the number of features required for sorting to just two.

the measured or received signal along specific directions in an iterative manner (as specified by the **F** matrix), which will maximize the measure of non-Gaussianity. Different measures of non-Gaussianity have been used to determine the optimal matrix **F**, such as the information-maximization approach[36] or methods based on cumulants[37] and negentropy.[38]

Using kurtosis (or the fourth-order cumulant) of a signal as a measure of its non-Gaussianity, an implementation of ICA originally derived by Hyvarinen and Oja[39] is discussed here. For the kth source signal s_k, the kurtosis is defined as

$$\text{Kurt}(s_k) = E\{s_k^4\} - 3[E\{s_k^2\}]^2 \tag{18.12}$$

where $E\{\}$ denotes the mathematical expectation. The kurtosis is negative for source signals whose amplitude has sub-Gaussian probability densities (probability distributions flatter than Gaussian), positive for super-Gaussian (probability distributions sharper than Gaussian), and zero for Gaussian densities. Maximizing the norm of the kurtosis leads to identification of the independent sources. This method was used in Ref. 33 to eliminate ECG interference in EEG signals.

A brief outline is presented here. For a more detailed presentation, the reader is referred to an excellent review.[31] The first step in ICA processing is to center **r** by subtracting its mean vector **m** so as to make **r** a zero-mean variable. After estimating the mixing matrix with centered data, the procedure is completed by adding the mean vector **m** of **r** back to the centered estimate that was subtracted in the first step. Subsequent to centering, the measured vector is sometimes whitened as part of preprocessing the signal before applying the ICA algorithm. Whitening the observed vector **r** (resulting in a flat power spectrum) removes the second-order relationships and produces a new vector $\mathbf{u} = \{u_1, u_2, \ldots, u_n\}$ whose components are uncorrelated. The variance of the individual components of **u** is unity. In matrix notation, this is

$$\mathbf{u} = \mathbf{Ur} \tag{18.13}$$

where **U** is the whitening matrix. Expressing **r** in terms of **s** and **A**, the model **u** becomes

$$\mathbf{u} = \mathbf{UAs} \tag{18.14}$$

The solution is

$$\mathbf{y} = \mathbf{Wu} \tag{18.15}$$

where **W** is the separating matrix for the measured signals after whitening (Fig. 18.6).

Maximizing the absolute value of the kurtosis of the components of **y** (Eq. 18.12), one of the columns of the separating matrix **W** is found, and so one independent component at a time is identified. The other columns are estimated subsequently. The algorithm has a cubic convergence and typically convergence by 20 iterations.[33] From Eqs. (18.13), (18.14), and (18.15), the output matrix of independent components **y** can be written as

$$\mathbf{y} = \mathbf{Wu} = \mathbf{WUr} = \mathbf{Fr} \tag{18.16}$$

The rows of this matrix are the time course of activation of the individual ICA components. A MATLAB implementation of the algorithm (FastICA) is available at http://www.cis.hut.fi/projects/ica/fastica/.[31]

APPLICATION *ICA has been used to separate ECG interference in EEG signals recorded from adult rats undergoing controlled ischemic brain injury and subsequent recovery.[33] The measured signal vector* **r** *consisted of two EEG signals recorded from right and left parietal cortex areas and one channel of ECG, as illustrated in Fig. 18.7. The signals shown at the top in Fig. 18.7 were recorded right after an asphyxic injury to the brain during which the EEGs became isoelectric. The early recovery from such injury is characterized by low-amplitude waveforms in the EEG that carry information about the early recovery mechanisms in the cortex. However, they are prone to corruption from the larger-amplitude ECG signals. The reconstructed EEG is free from any interference*

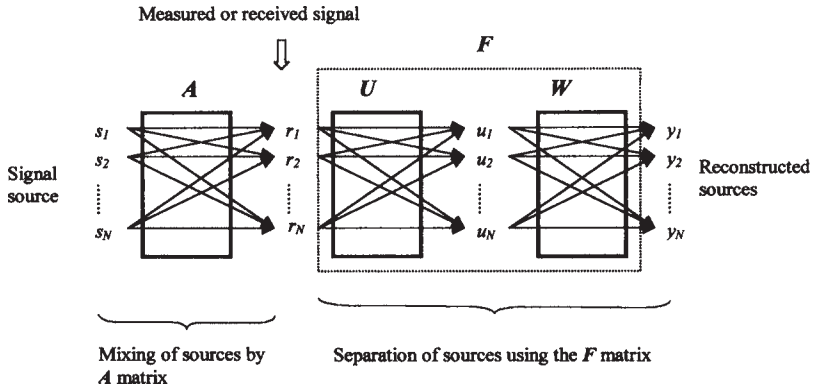

FIGURE 18.6 Schematic of independent components analysis technique for a blind source separation application. The signal from the sources gets mixed by an unknown mixing matrix **A** before being measured. The goal of ICA is to design a separating matrix **F** of linear spatial filters that will invert the measured signal vector **r** back into the independent components represented by the vector **y**. The components of **y** should be approximations of the source vector s although they (the reconstructed sources) do not have to match the order in which they appear in **s**.

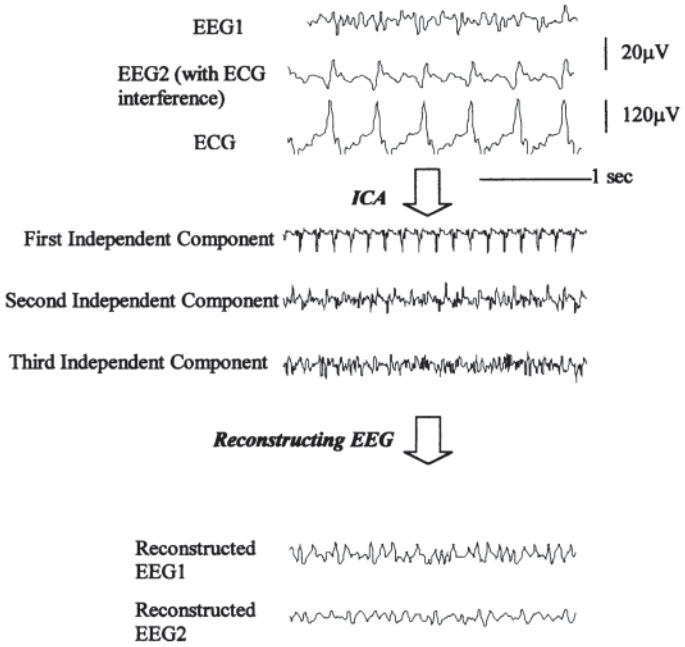

FIGURE 18.7 Independent components analysis has been used to separate ECG interference in the EEG.[33] Two channels of EEG and an interfering ECG waveform are shown at the top. The EEG data were recorded from an adult rat just after an episode of brain injury due to oxygen deprivation resulting in electrical silence in the brain. EEG during the early recovery stages shows very low-amplitude activity resulting in pronounced interference from the ECG waveforms. The independent components are used for reconstructing the two EEG waveforms shown at the bottom.

from ECG. The EEG was reconstructed from the independent components by setting the independent component in $\mathbf{y}$ *that corresponds to ECG to zero and evaluating* $\mathbf{r} = \mathbf{F}^{-1}\mathbf{y}$ *[from Eq. (18.16)].*

18.6 CROSS-CORRELATION AND COHERENCE ANALYSIS

Signal Types—Pairs of Stochastic Signals, Pairs of Stochastic and a Deterministic Signal, and Multichannel Signal Analysis. The signal-analysis methods discussed so far represented the measured signals along a set of orthogonal basis functions (Fourier, wavelet, or principal components) for noise removal (thus enhancing the signal-to-noise ratio), feature extraction, signal-source separation, etc. A bioengineer will also find situations where he or she has to deal with pairs of signals. Cross-correlation and coherence analysis are often simple tools that are extremely effective in quantifying relationships among pairs of biosignals.[40–46] Some of the typical application situations are as follows:

1. Detecting a known (deterministic) signal in a noisy environment
2. Estimating time delay or propagation delay between two biosignals (deterministic or stochastic signals in noise)
3. Estimating the transfer function of the signal-generating system (deconvolution, estimating the degree of linearity, etc.)

The cross-correlation between two measured biosignals $x(n)$ and $y(n)$ is defined statistically as $R_{yx}(k) = E[y(n)x(n + k)]$, where the operator E represents the statistical expectation or mean and k is the amount of time signal $x(n)$ is delayed with respect to $y(n)$. Given two time sequences of $x(n)$ and $y(n)$ each of N points, the commonly used estimate of cross-covariance $c_{yx}(k)$ is as follows[4]:

$$c_{yx}(k) = \begin{cases} \dfrac{1}{N}\displaystyle\sum_{n=0}^{N-k-1} [y(n) - \hat{m}_y][x(n + k) - \hat{m}_x] & k \geq 0 \\ \dfrac{1}{N}\displaystyle\sum_{n=0}^{N-k-1} [y(n + k) - \hat{m}_y][x(n) - \hat{m}_x] & k \leq 0 \end{cases} \tag{18.17}$$

where $\hat{m}_y$ and $\hat{m}_x$ are the sample means of the signal sequences $y(n)$ and $x(n)$, respectively. The sample means can be omitted in this expressions if the signals have been detrended using a MATLAB function such as *detrend*.

Statistics of the Estimate. The cross-correlation estimator is however biased[4] because

$$E[c_{yx}(k)] = \left(1 - \frac{|k|}{N}\right) C_{yx}(k) \tag{18.18}$$

where $C_{yx}(k)$ is the true value of the cross-covariance or cross-correlation at a lag of k. The bias reduces for increasing length N of the signal sequence. However, the variance of this estimate is complicated. If the two sequences are Gaussian and completely uncorrelated, then the variance in the estimate of the cross-correlation is $\sigma_x^2\sigma_y^2/N$, where σ_x^2 and σ_y^2 are variances of $x(n)$ and $y(n)$ (the white noise sequences), respectively. In general, estimation of correlation between two sequences has to be done cautiously, as outlined in Ref. 4. Each individual signal sequence is modeled using a parametric model outlined in Sec. 18.3.3. The cross-correlation between the two error sequences (should be white Gaussian noise if the model is accurate) that were generated by modeling each signal must be estimated. The 95 percent confidence limits on the cross-correlation between the error sequences must be computed using the variance expression given earlier. Finally, the cross-corre-lation between the

two signal sequences $x(n)$ and $y(n)$ can be estimated. Values that exceed the 95 percent confidence limits of the cross-correlation between the two error sequences indicate statistically significant correlation.

18.6.1 Detection of Known Signals (Deterministic) in White Gaussian Noise

The simple cross-correlation estimator is used extensively in the form of a matched filter implementation to detect a finite number of known signals (in other words, simultaneous acquisition of multiple channels of known signals). When these deterministic signals are embedded in white Gaussian noise, the matched filter (obtained from cross-correlation estimate at zero lag, $k = 0$, between the known signal sequence and the observed noisy signal sequence) gives the optimum detection performance (in the Bayesí sense[29]).

If the known N-dimensional signal vector is $\mathbf{S}$ (could be any signal of fixed morphology, such as the ECG, evoked potential, action potentials, etc.) and the measured N-dimensional observation vector with white Gaussian noise is $\mathbf{X}$, then we evaluate a likelihood ratio that is defined as

$$\sum_{i=1}^{N} \frac{S_i X_i}{\|\mathbf{S}\|^2}$$

and compare it against a Bayes threshold Λ_0 identical to the one described in the section on principal component analysis. When the likelihood ratio exceeds the threshold, the detector flags the presence of the known signal in the measurement.

APPLICATION *Matched filters can be used to detect and sort action potentials in noise. The multiunit data in Fig. 18.8b were collected from the ventral posteriolateral thalamic nuclei using a microelectrode. The segment shows two different types of action potentials (corresponding to two*

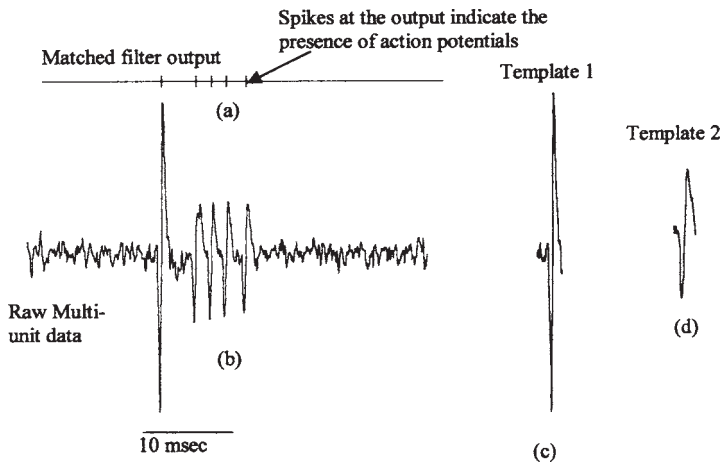

FIGURE 18.8 Cross-correlation-based method to detect the presence of spikes or action potentials in multiunit activity. The continuous waveform with the action potentials is shown in (*b*). The shapes of the action potentials to be detected are known a priori (deterministic signal). Using a cross-correlation of the known signal [templates (*c*) and (*d*)] with the measured signal, the presence of action potentials is indicated by a vertical line in (*a*). This approach is optimal in the presence of white Gaussian noise.

different neurons at different locations relative to the microelectrode). Action potential shapes from trial data were recognized and stored as templates. Two of the templates presumably corresponding to two different neurons are shown in Fig. 18.8c and d. Action potentials (or spikes) in the subsequent data set are detected (indicated by vertical lines in the detector output in Fig. 18.8a) and sorted based on these templates (detecting a known signal in noise). The noise is assumed to be white Gaussian, and the matched filter works reasonably well in detecting these spikes. However, the detector often fails when two action potentials overlap.

In contrast to the principal component (PCA)ñbased approach for detection and discrimination, the cross-correlation or matched-filter approach is based on a priori knowledge of the shape of the deterministic signal to be detected. However, the PCA-based method also requires some initial data (although the data could be noisy, and the detector does not need to know a priori the label or the class of the different signals) to evaluate the sample covariance matrix **K** and its eigenvectors. In this sense, the PCA-based detector operates in an unsupervised mode. Further, the matched-filter approach is optimal only when the interfering noise is white Gaussian. When the noise is colored, the PCA-based approach will be preferred.

18.6.2 Coherence Estimation

The estimation of coherence is useful to determine if *spectral* components between two signal sequences are significantly correlated. It is also useful in determining the linearity of biological systems. Although most biological systems can be expected to be nonlinear, changes in the degree of nonlinearity often are correlated with changes in underlying physiology or pathology. For instance, Kong et al.[47] used coherence to monitor changes in blood oxygenation level in the brain by calculating the linearity of the somatosensory evoked responses. Assuming that $x(n)$ and $y(n)$ are the input and output of a linear system, the cross-spectral density $S_{yx}(f)$ is given by $S_{yx}(f) = S_x(f)H(f)$, where $S_x(f)$ is the amplitude spectral density of the input signal at f and $H(f)$ is the transfer function of the linear system. A coherence function defined as

$$\hat{Coh}_{xy}(f) = \frac{|S_{yx}(f)|^2}{S_y(f)S_x(f)} \tag{18.19}$$

is therefore unity (maximum value) for linear systems. Consequently, any deviation from unity can be used as an indicator of the presence of nonlinearity in the system. The coherence function is estimated using the values of $S_x(f)$, $S_y(f)$, and $S_{yx}(f)$ obtained through FFTs. In MATLAB, the *spectrum* and *cohere* functions can be used to evaluate the cross-spectral density and the coherence.

Statistics of the Estimates. The mean and the variance of the sample estimates of the coherence function are derived in Ref. 4, and I only reproduce the final results here. In general, the cross-spectral density is evaluated by doing the Fourier transform of a windowed (using a lag window) sequence of cross-correlation estimates. The choice of the smoothing window therefore determines the variance in the estimates of cross-spectral density (numerator in the expression of coherence). The variance of the smoothed coherence estimator is given by[4]

$$\frac{1}{2V}[1 - \hat{Coh}_{xy}(f)]^2 \tag{18.20}$$

where V is a variance reduction factor that depends on the smoothing window used. The different types of windows and their respective variance reduction factors can be found in Ref. 4. In general, the user has to be careful in choosing a large enough lag window that it includes all the largest

cross-correlation estimates. An insufficient length in the lag window can give unreasonably large bias in the estimate of cross-spectral density. The coherence estimate is also biased because

$$E[\hat{Coh}_{xy}(f)] \approx \left(1 - \frac{\tau}{NT}\right) Coh_{xy}(f) \tag{18.21}$$

where τ is the time shift between the two sequences[4,48] that can be found by locating the lag corresponding to the peak estimate of cross-correlation. The source of bias can then be minimized by aligning the signals so that τ becomes negligibly small before estimating the cross-spectral density.

APPLICATION *An adaptive implementation of coherence estimation has been used in Ref. 47 to track hypoxic injury related changes in somatosensory evoked potential signals, as shown in Fig. 18.9. Fig. 18.9a shows the change in amplitude of one of the peaks in the somatosensory evoked response (SEP amplitude) from a cat at different levels of oxygen concentration in the inhaled air. After the administration of 9 percent oxygen, the SEP amplitude shows a gradual decrease until the oxygen is restored to 100 percent. A linearity index derived using the coherence measure gives a more sensitive indication of injury, as shown in Fig. 18.9b.*

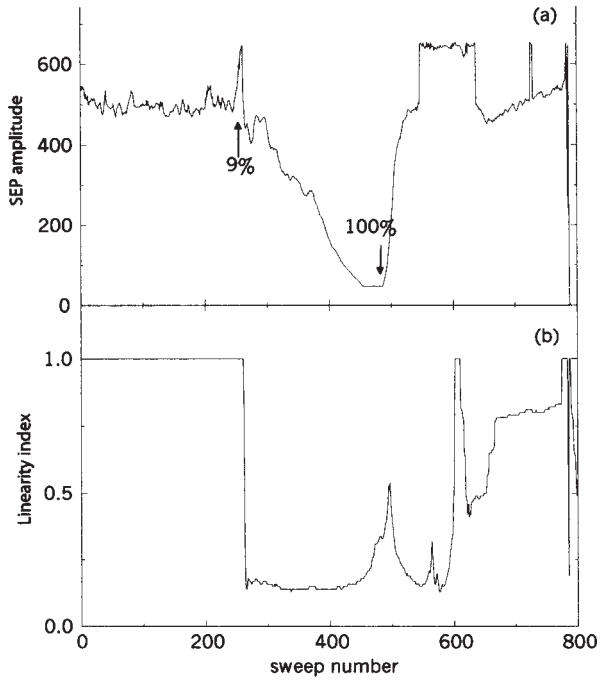

FIGURE 18.9 Adaptive coherence method to track changes in the somatosensory evoked potentials (SEPs).[47] A linearity index [shown in (*b*)] derived from the coherence of SEPs from the brain under low oxygen conditions with the SEPs derived during normal oxygen conditions was shown to be a very sensitive indicator of changes in brain oxygen content. The conventional amplitude of the SEP [shown in (*a*)] was not found to be that sensitive.

18.7 CHAOTIC SIGNALS AND FRACTAL PROCESSES

Signal Typesó Random Signals That Are Not Stochastic But Are Generated by a Nonlinear Dynamic System. Fractal signals are scale-invariant, meaning that they look very similar at all levels of magnification. There is very good evidence to indicate that the beat-to-beat interval in heart rate is a fractal. Chaotic signals, on the other hand, are deterministic signals that cannot be predicted beyond a short time in the future.[2] They are extremely sensitive to initial conditions. Further, their behavior does not depend on any random inputs. The nonlinear time-series analysis techniques discussed briefly here have been used to analyze heart rate,[49ñ51] nerve activity,[52] renal flow,[53] arterial pressure, EEG,[54,55] and respiratory waveforms.[56]

The randomness of a chaotic time series is not due to noise but rather due to deterministic dynamics of a small number of dynamical variables in the nonlinear generating system. Therefore, they need to be distinguished from stochastic processes discussed earlier and analyzed appropriately. The randomness in a chaotic signal is therefore not revealed by statistical analysis but by dynamical analysis based on phase-space reconstruction. The phase space of a dynamical system (both nonlinear and linear) is the space spanned by its dynamical variables, and the phase plot is the plot of the time variation in the dynamical variables within the phase space. The phase plot of a dynamical system generating a chaotic time series is a strange attractor whose dimensionality (dimension of the set of points comprising the phase plot) is not an integer but a fraction. Hence they are said to have a *fractal* dimension. Standard objects in Euclidean geometry have integer dimensions. For instance, the dimension of a point in space is zero, that of a line is one, and that of an area is two. The fractal dimension D of a strange attractor (name of the phase plot of a nonlinear system generating a chaotic signal) is related to the minimum number of dynamical variables needed to model the dynamics of the strange attractor. Analysis of chaotic signals typically is geared toward (1) understanding how complex the nonlinear system is, (2) determining if it is chaotic or not, (3) determining the number of dynamic variables that dominate the system, and (4) assessing the changes in dynamic behavior of a system with different rhythms.

The following subsection details some methods to find out if a given random time series from a biological source is chaotic. Further, methods to determine the dimensionality and draw the phase plot or portrait of the chaotic signal are outlined briefly.

18.7.1 Analysis of Chaotic Signals

Frequency Analysis. The statistical analysis of chaotic signals includes spectral analysis to confirm the absence of any spectral lines, since chaotic signals do not have any periodic deterministic component. Absence of spectral lines would indicate that the signal is either chaotic or stochastic. However, chaos is a complicated nonperiodic motion distinct from stochastic processes in that the amplitude of the high-frequency spectrum shows an exponential decline. The frequency spectrum can be evaluated using FFT-based methods outlined the earlier sections.

Estimating Correlation Dimension of a Chaotic Signal. One of the ways to measure the dimensionality of the phase portrait of a chaotic signal is through what is known as the *correlation dimension* D_2. It is defined as

$$D_2 = \frac{\log C(r)}{\log r}\dagger$$

as $r \to 0$, where r is the length of the side of the hypercubes that are needed to cover the phase plot and $C(r)$ is correlation sum. The correlation sum is defined as $C(r) = \sum_{i=1}^{M(r)} p_i^2$, where p_i is the probability of a finding a single point belonging to the phase plot within the hypercube. $M(r)$ is the number of m-dimensional cells or hypercubes of side r needed to cover the entire phase plot.

One of the key steps in the analysis is therefore to reconstruct the phase plot or phase space of the dynamical system. The embedding theorems of Takens[57] and Sauer et al.[58,59] help us in reconstructing the phase space using a time series from the dynamical system rather than using the dynamical variables of the system. Given a discrete time series $x(n)$ from a biological source, we construct a k-dimensional point in the phase space using k - 1 time-delayed samples of the same series as represented by

$$\mathbf{x}(n) = \{x(n), x(n + \tau), x(n + 2\tau), \ldots, x[n + (k - 1)\tau]\} \quad \dagger$$

The selection of the time delay τ is done in such a way that it makes every component of phase space uncorrelated. Therefore, τ is determined from estimate of the autocorrelation function of the time series. The time lag that corresponds to the first zero in the autocorrelation is often chosen as a good approximation for τ.[60,61] The determination of D_2 in practice can be done using the Grassberger-Procaccia algorithm[62] outlined below. Consider a pair of points in space with m dimensions ($m < k$) at time instants i and j:

$$\mathbf{x}_m(i):\{x(i), x(i + \tau), x(i + 2\tau), \ldots, x[i + (m - 1)\tau]\}$$
$$\mathbf{x}_m(j):\{x(j), x(j + \tau), x(j + 2\tau), \ldots, x[j + (m - 1)\tau]\} \quad \dagger$$

where m is called the *embedding dimension* of the phase plot. However, m is not known a priori. Therefore, we determine the correlation dimension D_2 for different embedding dimensions of the attractor in the phase space. The minimum embedding dimension is then given by m + 1, where m is the embedding dimension above which the measured value of the correlation dimension D_2 for the corresponding phase plot remains constant. The Euclidean distance between the two points is given by $r_{ij}(m) = \|\mathbf{x}_m(i) - \mathbf{x}_m(j)\|$. For a critical distance r, a correlation integral (an approximation of the correlation sum defined earlier) is evaluated that gives the probability of the distance between the two given points being less than r:

$$C_2(r, m) = \frac{1}{N^2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \theta[r - \|\mathbf{x}_m(i) - \mathbf{x}_m(j)\|] \quad \dagger$$

where†† $N = k - (m - 1)\tau$
θ = Heaviside function
$C_2(r, m)$ = correlation integral

An estimate of the correlation dimension D_2 is given by

$$D_2(m) = \frac{\log[C_2(r, m)]}{\log(r)}$$

The log-log plot of $C_2(r, m)$ versus r corresponding to the given m has a linear region called the *scaling region*, the slope of which gives an estimate of the correlation dimension. The reliability of the estimated slope in the linear scaling region can be a major source of error in the Grassberger-Procaccia algorithm. If N_c point pairs $\{\log[C_2(r_i, m)], \log(r_i) | i = 1, 2, \ldots, N_c\}$ exist in the scaling region, then $D_2(m)$ is given by[63]

$$D_2(m) = \frac{N_c \sum_{i=1}^{N_c} \log(r_i) \log[C_2(r_i, m)] - \sum_{i=1}^{N_c} \log(r_i) \sum_{i=1}^{N_c} \log[C_2(r_i, m)]}{N_c \sum_{i=1}^{N_c} [\log(r_i)]^2 - \left[\sum_{i=1}^{N_c} \log(r_i)\right]^2} \quad \dagger$$

The value of m beyond which $D_2(m)$ gradually saturates determines the embedding dimension m_c. That is, $D_2(m_c) = D_2(m_c + 1) = D_2(m_c + 2) = \ldots = D_2(m_c)$ gives an estimate of the correlation dimension.

Phase-Space Analysis. Having determined the time lag t for decorrelating the components of the phase space and the embedding dimension, a line joining the m_c dimensional points given by

$$\mathbf{x}_{m_c}(j):\{x(j), x(j+\tau), x(j+2\tau), \ldots, x[j+(m_c-1)\tau]\} \quad \dagger$$

(where j = 1, 2, Ö, n) gives a portrait of the evolution of dynamical variables of the system. According to the Takens theorem, for an attractor having an integer-dimension manifold, the phase plot obtained from the time series preserves the topological properties (such as dimension) of the attractor. Phase-plane analysis of chaotic processes enables quantitative description subsequently through calculation of Lyapunov exponents, Poincare mapping, etc. The goal of this characterization is to obtain a portrait of evolution of the state variables of the system. Additional parameters such as the Lyapunov exponents and complexity measures can be determined as well once the biological process begins to exhibit chaotic behavior. The user is referred to Refs. 64 to 66 for further information.

There are several issues that a user has to be aware of while doing a nonlinear time-series analysis. If the biosignal in question is indeed generated by a deterministic dynamical system, then the sampling frequency of the time series should be sufficiently large to capture the deterministic rule governing the evolution of the series. The length of the sequence should also be sufficiently large. Fortunately, there is a rule of thumb given by Ruelle[67] for the minimum length of time series that is quite helpful. It states that estimates of the dimension of the phase plot $D \geqslant 2 \log_{10} N$ should be regarded as unreliable where N is the length of the time series. Finally, noise contamination during measurement could obscure detection of deterministic dynamics and hence degrade forecasting of biologically significant events. The user is referred to Ref. 66 for MATLAB codes for the Grassberger-Procaccia algorithm, phase-space reconstruction, and forecasting.

Surrogate Data to Compare Chaotic and Stochastic Processes. Surrogate data are used in nonlinear time-series analysis to make useful comparisons between the given biosignal, which could be chaotic, and a set of artificially generated stochastic signals that share the same essential statistical properties (mean, variance, and the power spectrum) with the biosignal.[66,68] If the measured topological properties of the biosignal lie within the standard deviation of the topological properties measured from the artificially generated surrogate data, then the null hypothesis that the biosignal is just random noise cannot be ruled out. A systematic method for generating surrogate data with the same mean, variance, and power spectrum as a given biosignal is detailed in Ref. 68, and a MATLAB code is available in Ref. 66. The main steps in the procedure are outlined below.

The DFT $X(k)$ of the observed signal sequence $x(n)$ is evaluated using Eq. (18.1). Keeping the magnitudes of the Fourier coefficients intact, the phases of the Fourier coefficients are randomized (maintaining symmetry about the midpoint so that we still get a real time series after inverse Fourier transform). The Fourier sequence is now inverted to produce a real time series that is a surrogate of the original data. With different random assignments for the phase values of Fourier coefficients, a different time series can be generated every iteration. All the different time series generated using this procedure would be part of the surrogate data to the biosignal. The random phase assignments are done as follows: If the length of the measured sequence is N (assumed to be even), a set of random phase values $\phi_m \epsilon [0, \pi]$, for m = 2, 3, Ö, $N/2$, is generated using a function such as *rand* in MATLAB. The phase of the Fourier coefficients are randomized to generate a new sequence of Fourier coefficients $X_s(k)$ as shown below:

$$X_s(k) = \begin{cases} X(k) & \text{for } k = 1 \text{ and } \dfrac{N}{2}+1 \\ |X(k)|e^{j\phi_m} & \text{for } k = 2, 3, \ldots, \dfrac{N}{2} \\ |X(N-k+2)|e^{j\phi_{N-k+2}} & \text{for } k = \dfrac{N}{2}+2, \dfrac{N}{2}+3, \ldots, N \end{cases} \quad \dagger$$

For every new set of $\phi_m \epsilon [0, \pi]$, we can generate a new surrogate time series $x_s(n)$ by performing the inverse Fourier transform of $X_s(k)$.

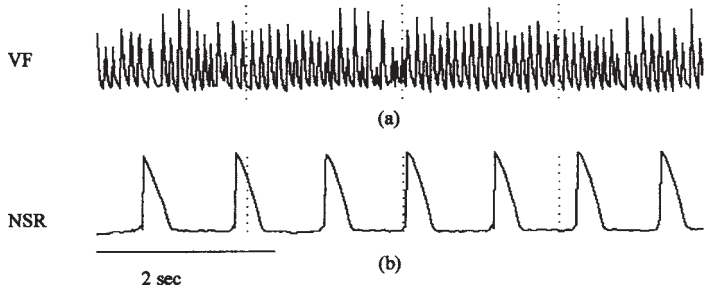

FIGURE 18.10 Action potential waveforms from an isolated heart (Langendorff preparation) are shown[63] (*a*) during ventricular fibrillation (VF) and (*b*) during normal sinus rhythm (NSR).

APPLICATION *Myocardial cell action potentials collected from a Langendorff setup of an isolated rabbit heart experiment have been found to exhibit chaotic behavior.[63] Two different cardiac cell action potential waveforms are illustrated in Fig. 18.10 corresponding to ventricular fibrillation (VF) in (a) and normal sinus rhythm (NSR) in (b). The corresponding frequency spectrum is shown in Fig. 18.11. VF certainly shows chaotic characteristics in Fig. 18.11a, whereas NSR shows distinct spectral lines in Fig. 18.11b, indicating the presence of deterministic components. Using the calculated correlation dimensions for the two different action potentials (VF = 5.629, NSR = 2.704), the embedding dimension is m = 6 for VF. The projections of the*

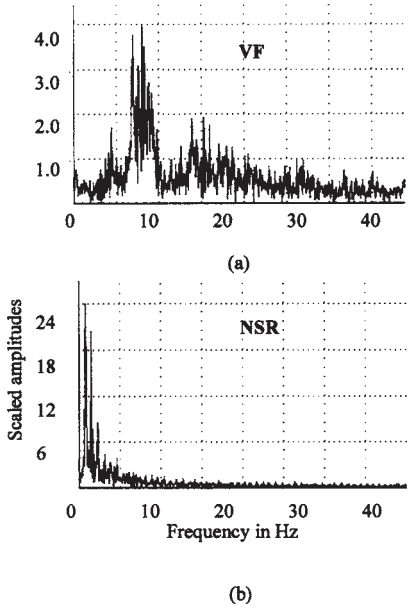

FIGURE 18.11 Power spectrum of the action potential waveform during[63] (*a*) VF and (*b*) NSR. The power spectrum of VF has no spectral components or peaks in contrast with the power spectrum of NSR, which has distinct spectral lines indicating that it is a periodic and not a stochastic or chaotic waveform.

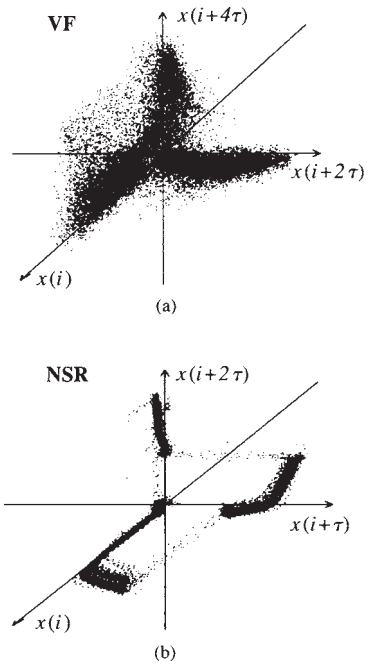

FIGURE 18.12 Phase plots of the action potential waveform during[63] (*a*) VF and (*b*) NSR. The time delay required to make the individual components of the phase plot uncorrelated was determined to be $6\delta t$, and $41\delta t$, respectively, where $\delta t = 5$ ms is the sampling interval. The embedding dimension of the VF waveform was determined to be 6, and that of the NSR waveform was 3. The phase plot of the VF waveform is shown in (*a*) spanned by $\{x(i), x(i + 2\tau), x(i + 4\tau)\}$, demonstrating a strange attractor, whereas the phase plot of the NSR waveform is shown in (*b*) spanned by $\{x(i), x(i + \tau), x(i + 2\tau)\}$.

trajectories for VF in a three-dimensional phase space spanned by {x(i), x(i + 2τ), x(i + 4τ)} is shown in Fig. 18.12a, whereas for NSR the embedding dimension m = 3. The corresponding three-dimensional phase portrait spanned by {x(i), x(i + τ), x(i + 2τ)} is shown in Fig. 18.12b. For VF and NSR, the decorrelating time τ was determined to be 6δt, and 41δt respectively, where δt = 5 ms is the sampling interval. Obviously, VF demonstrates chaotic character, whereas NSR demonstrates a periodic or quasi-periodic nature.

REFERENCES

1. M. Akay, *Biomedical Signal Processing*, Academic Press, San Diego, 1994.
2. E. N. Bruce, *Biomedical Signal Processing and Signal Modeling*, Wiley, New York, 2001.
3. S. M. Kay and S. L. Marple, Spectrum analysis: A modern perspective, *Proc. IEEE* **69**(11):1380ñ1419 (1981).
4. R. Shiavi, *Introduction to Applied Statistical Signal Analysis*, 2d ed. Academic Press, San Diego, 1999.
5. P. M. T. Broersen, Facts and fiction in spectral analysis, *IEEE Trans. Instum. Measur.* **49**:766ñ772 (2000).

6. P. M. T. Broersen, Finite sample criteria for autoregressive order selection, *IEEE Trans Signal Proc.* **48**:3550ñ3558 (2000).
7. J. Muthuswamy and N. V. Thakor, Spectral analysis of neurological signals, *J. Neurosci. Methods* **83**:1ñ14 (1998).
8. H. Akaike, A new look at the statistical model identification, *IEEE Trans. Automatic Control* **AC-19**:716ñ723 (1974).
9. J. P. Burg, *Maximum entropy spectral analysis*, Dept. of Geophysics, Stanford University, Stanford, Calif., 1975.
10. K. A. Wear, R. F. Wagner, and B. S. Garra, A comparison of autoregressive spectral estimation algorithms and order determination methods in ultrasonic tissue characterization, *IEEE Trans Ultrason. Ferroelect. Frequency Control* **42**:709ñ716 (1995).
11. F. Hlawatsch and G. F. Boudreaux-Bartels, Linear and quadratic time-frequency signal representations, *IEEE Signal Processing Magazine*, pp. 21ñ67, 1992.
12. J. Z. Wang, Wavelets and imaging informatics: A review of the literature, *J. Biomed. Inform.* **34**:129ñ141 (2001).
13. V. von Tscharner and K. R. Thulborn, Specified-resolution wavelet analysis of activation patterns from BOLD contrast fMRI, *IEEE Trans. Med. Imaging* **20**:704ñ714 (2001).
14. M. Desco, J. A. Hernandez, A. Santos, and M. Brammer, Multiresolution analysis in fMRI: Sensitivity and specificity in the detection of brain activation, *Hum. Brain Mapp.* **14**:16ñ27 (2001).
15. D. F. Schomer, A. A. Elekes, J. D. Hazle, J. C. Huffman, S. K. Thompson, C. K. Chui, and W. A. Murphy Jr.,, Introduction to wavelet-based compression of medical images, *Radiographics* **18**:469ñ481 (1998).
16. K. Wang, H. Begleiter, and B. Porjesz, Spatial enhancement of event-related potentials using multiresolution analysis, *Brain Topogr.* **10**:191ñ200 (1998).
17. P. J. Phillips, The design of matching pursuit filters, *Network* **9**:1ñ17 (1998).
18. K. L. Park, K. J. Lee, and H. R. Yoon, Application of a wavelet adaptive filter to minimise distortion of the ST-segment, *Med. Biol. Eng. Comput.* **36**:581ñ586 (1998).
19. N. D. Panagiotacopulos, J. S. Lee, M. H. Pope, and K. Friesen, Evaluation of EMG signals from rehabilitated patients with lower back pain using wavelets, *J. Electromyogr. Kinesiol.* **8**:269ñ278 (1998).
20. J. Jalife, O. Berenfeld, A. Skanes, and R. Mandapati, Mechanisms of atrial fibrillation: Mother rotors or multiple daughter wavelets, or both? *J. Cardiovasc. Electrophysiol.* **9**:S2ñ12 (1998).
21. R. Gobbele, H. Buchner, and G. Curio, High-frequency (600 Hz) SEP activities originating in the subcortical and cortical human somatosensory system, *Electroencephalogr. Clin. Neurophysiol.* **108**:182ñ189 (1998).
22. Q. Fang and I. Cosic, Protein structure analysis using the resonant recognition model and wavelet transforms, *Australas. Phys. Eng. Sci. Med.* **21**:179ñ185 (1998).
23. P. Carre, H. Leman, C. Fernandez, and C. Marque, Denoising of the uterine EHG by an undecimated wavelet transform, *IEEE Trans. Biomed. Eng.* **45**:1104ñ1113 (1998).
24. N. Aydin, S. Padayachee, and H. S. Markus, The use of the wavelet transform to describe embolic signals, *Ultrasound Med. Biol.* **25**:953ñ958 (1999).
25. V. J. Samar, A. Bopardikar, R. Rao, and K. Swartz, Wavelet analysis of neuroelectric waveforms: A conceptual tutorial, *Brain Lang.* **66**:7ñ60 (1999).
26. M. Akay and C. Mello, Wavelets for biomedical signal processing, presented at 19th International Conference of IEEE/EMBS, Chicago, 1997.
27. R. Kronland-Martinet, J. Morlet, and A. Grossmann, Analysis of sound patterns through wavelet transforms, *Int. J. Pattern Recog. Artif. Intell.* **1**:273ñ301 (1987).
28. B. Gramatikov, S. Yi-chun, P. Caminal, H. Rix, and N. V. Thakor, Multiresolution wavelet analysis of the body surface ECG before and after angioplasty, *Ann. Biomed. Eng.* **23**:553ñ561 (1995).
29. H. L. van Trees, *Detection, Estimation and Modulation Theory*, Vol. **I**, Wiley, New York, 1969.
30. R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, Wiley, New York, 2001.
31. A. Hyvarinen and E. Oja, Independent component analysis: Algorithms and applications, *Neural Networks* **13**:411ñ430 (2000).
32. S. Makeig, A. Bell, T. P. Jung, and T. J. Sejnowski, Independent component analysis of electroencephalographic data, in *Advances in Neural Information Processing Systems*, Vol. **8**, pp. 145ñ151, MIT Press, Cambridge, Mass., 1996.
33. S. Tong, A. Bezerianos, J. Paul, Y. Zhu, and N. Thakor, Removal of ECG interference from the EEG recordings in animals using independent component analysis, *J. Neurosci. Methods* **108**:11ñ17 (2001).

34. R. Vigario, J. Sarela, and E. Oja, Independent component analysis in wave decomposition of auditory evoked fields, presented at International Conference on Artificial Neural Networks (ICANN í98), Skovde, Sweden, 1998.
35. R. Vigario, V. Jousmaki, M. Hamalainen, R. Hari, and E. Oja, Independent component analysis for identification of artifacts in magnetoencephalographic recordings, in *Advances in Neural Information Processing Systems*, pp. 229ñ235, MIT Press, Cambridge, Mass., 1998.
36. J. P. Nadal and N. Parga, Nonlinear neurons in the low noise limit: a factorial code maximized information transfer, *Network* **5**:565ñ581 (1994).
37. P. Comon, Independent component analysis: A new concept? *Signal Processing* **36**:287ñ314 (1994).
38. M. Girolami and C. Fyfe, Extraction of independent signal sources using a deflationary exploratory projection pursuit network with lateral inhibition, *IEEE Proc. Vision Image Signal Proc.* **14**:299ñ306 (1997).
39. A. Hyvarinen and E. Oja, A fast fixed-point algorithm for independent component analysis, *Neural Comput.* **9**:1483ñ1492 (1997).
40. P. H. Boeijinga and F. H. Lopes da Silva, A new method to estimate time delays between EEG signals applied to beta activity of the olfactory cortical areas, *Electroencephalogr. Clin. Neurophysiol.* **73**:198ñ205 (1989).
41. Z. Bohdanecky, P. Lansky, and T. Radil, An integral measure of the coherence function between pairs of EEG recordings, *Electroencephalogr. Clin. Neurophysiol.* **54**:587ñ590 (1982).
42. P. Brown, A. Oliviero, P. Mazzone, A. Insola, P. Tonali, and V. Di Lazzaro, Dopamine dependency of oscillations between subthalamic nucleus and pallidum in Parkinsonís disease, *J. Neurosci.* **21**:1033ñ1038 (2001).
43. S. Cerutti, M. Alberti, G. Baselli, O. Rimoldi, A. Malliani, M. Merri, and M. Pagani, Automatic assessment of the interaction between respiration and heart rate variability signal, *Med. Prog. Technol.* **14**:7ñ19 (1988).
44. M. P. Davey, J. D. Victor, and N. D. Schiff, Power spectra and coherence in the EEG of a vegetative patient with severe asymmetric brain damage, *Clin. Neurophysiol.* **111**:1949ñ1954 (2000).
45. L. Fendelander, P. W. Hsia, and R. J. Damiano, Jr., Spatial coherence: A new method of quantifying myocardial electrical organization using multichannel epicardial electrograms, *J. Electrocardiol.* **30**:9ñ19 (1997).
46. T. B. Kuo, C. C. Yang, and S. H. Chan, Transfer function analysis of ventilatory influence on systemic arterial pressure in the rat, *Am. J. Physiol.* **271**:H2108ñ2115 (1996).
47. X. Kong and N. Thakor, Adaptive coherence estimation reveals nonlinear processes in injured brain, in *IEEE Proc. Intl. Conf. Acoustics, Speech, and Signal Processing, ICASSP-93*, **1**:89ñ90 (1993).
48. G. Carter, Coherence and time-delay estimation, in C. Chen (ed.), *Signal Processing Handbook*, Dekker, New York, 1988.
49. M. R. Guevara, L. Glass, and A. Shrier, Phase locking, period doubling bifurcations, and irregular dynamics in periodically stimulated cardiac cells, *Science* **214**:1350ñ1352 (1981).
50. D. R. Chialvo, D. C. Michaels, and J. Jalife, Supernormal excitability as a mechanism of chaotic dynamics of activation in cardiac purkinje fibers, *Circ. Res.* **66**:525ñ545 (1990).
51. K. M. Stein, N. Lippman, and P. Kligfield, Fractal rhythms of the heart, *J. Electrocardiol.* **24**(suppl.):72ñ76 (1992).
52. Z. S. Huang, G. L. Gebber, S. Zhong, and S. Barman, Forced oscillations in sympathetic nerve discharge, *Am. J. Physiol.* **263**:R564ñ571 (1992).
53. K. P. Yip, N. H. Holstein-Rathlou, and D. J. Marsh, Chaos in blood flow control in genetic and renovascular hypertensive rats, *Am. J. Physiol.* **261**:F400ñ408 (1991).
54. L. D. Iasemidis, L. D. Olson, R. S. Savit, and J. C. Sackellares, Time dependencies in the occurrences of epileptic seizures, *Epilepsy Res.* **17**:81ñ94 (1994).
55. C. A. Skarda and W. J. Freeman, How brains make chaos in order to make sense of the world, *Behav. Brain Sci.* **10**:161ñ195 (1987).
56. D. Hoyer, K. Schmidt, R. Bauer, U. Zwiener, M. Kohler, B. Luthke, and M. Eiselt, Nonlinear analysis of heart rate and respiratory dynamics, *IEEE Eng. Med. Biol. Mag.* **16**:31ñ39 (1997).
57. F. Takens, Detecting strange attractors in turbulence, D. A. Rand and L. S. Young (eds.), in *Dynamical Systems and Turbulence*, Springer, Berlin, 1981.
58. T. Sauer and J. A. Yorke, Rigorous verification of trajectories for computer simulation of dynamical systems, *Nonlinearity* **4**:961ñ979 (1991).
59. T. Sauer, J. Yorke, and M. Casdagli, Embedology, *J. Stat. Phys.* **65**:579ñ616 (1994).

60. D. Holton and R. M. May, Distinguishing chaos from noise, in *The Nature of Chaos*, Chap. 7, Oxford University Press, Oxford, England, 1993.

61. A. M. Fraser and H. L. Swinney, Using mutual information to find independent coordinates of strange attractors, *Phys. Rev. [A]* **33**:1134ñ1139 (1986).

62. P. Grassberger and I. Procaccia, Characterization of strange attractors, *Phys. Rev. Lett.* **50**:346ñ349 (1983).

63. X. S. Zhang and Y. S. Zhu, The extraction of dynamical information of myocardial cell electrophysiology by fractal and chaotic analysis methods, *Automedica* **18**:9ñ26 (1999).

64. F. Kasper and H. G. Schuster, Easily calculable measure for the complexity of spatiotemporal patterns, *Phys. Rev. [A]* **36**:842ñ848 (1987).

65. A. Lempel and J. Ziv, On the complexity of finite sequences, *IEEE Trans. Inform. Technol.* **22**:75ñ81 (1976).

66. B. Henry, N. Lovell, and F. Camacho, Nonlinear dynamics time series analysis, M. Akay (ed.), in *Nonlinear Biomedical Signal Processing*, Vol. II, pp. 1ñ39, IEEE Press, New York, 2000.

67. D. Ruelle, Deterministic chaos: The science and the fiction, *Proc. R. Soc. Lond. [A]* **427**:241ñ248 (1990).

68. J. Theiler, B. Galdrikian, A. Longtin, S. Eubank, and J. D. Farmer, Using surrogate data to detect nonlinearity in time series, M. Casdagli and S. Eubank (eds.), in *Nonlinear Modeling and Forecasting*, Addison-Wesley, Reading, Mass., 1992.